

6 Protein Crystal Receptor
NOVEL HUMAN ~~MEMBRANE PROTEIN~~ AND
POLYNUCLEOTIDES ENCODING THE SAME

The present application claims the benefit of U.S. Provisional Application Number 60/175,764 which was filed on January 12, 2000 and is herein incorporated by reference in its entirety.

1. INTRODUCTION

The present invention relates to the discovery, identification and characterization of a novel human polynucleotide sequence encoding a membrane associated protein/receptor. The invention encompasses the described polynucleotides, host cell expression systems, the encoded protein, fusion proteins, polypeptides and peptides, antibodies to the encoded proteins and peptides, and genetically engineered animals that lack the disclosed sequence, or over express the disclosed sequence, or antagonists and agonists of the protein, and other compounds that modulate the expression or activity of the protein encoded by the disclosed sequence that can be used for diagnosis, drug screening, clinical trial monitoring, and/or the treatment of physiological or behavioral disorders.

2. BACKGROUND OF THE INVENTION

Membrane receptor proteins are integral components of the mechanism through which cells sense their surroundings, regulate cellular functions, and maintain physiological homeostasis. Accordingly, membrane receptor proteins are often involved in signal transduction pathways that control cell physiology, chemical communication, and gene expression. A particularly relevant class of membrane receptors are those typically characterized by the presence of 7 conserved transmembrane domains that are interconnected by nonconserved hydrophilic loops. Such, "7TM receptors" include a superfamily of receptors known as G-protein coupled receptors (GPCRs). GPCRs are typically

involved in signal transduction pathways involving G-proteins or PPG proteins. As such, the GPCR family includes many receptors that are known targets for therapeutic agents.

3. SUMMARY OF THE INVENTION

5 The present invention relates to the discovery, identification and characterization of nucleotides that encode a novel GPCR, and the corresponding novel GPCR amino acid sequence. The GPCR described for the first time herein, is a transmembrane proteins that spans the cellular membrane and is involved in
10 signal transduction after ligand binding. The described GPCR has structural motifs found in the 7TM receptor family. The novel human GPCR described herein, encodes a protein of 313 amino acids in length (see SEQ ID NO: 2). The described GPCR has an atypical hydrophobic leader sequence, the characteristic seven
15 transmembrane regions (of about 20-30 amino acids), as well as several predicted cytoplasmic domains.

 The invention encompasses the nucleotides presented in the Sequence Listing, host cells expressing such nucleotides, and the expression products of such nucleotides, and: (a) nucleotides
20 that encode mammalian homologues of the described novel GPCR (NGPCR) including the specifically described human NGPCR, and the human NGPCR gene product; (b) nucleotides that encode one or more portions of the NGPCR that correspond to functional domains, and the polypeptide products specified by such nucleotide sequences,
25 including but not limited to the novel regions of the extracellular domain(s) (ECD), one or more transmembrane domain(s) (TM) first disclosed herein, and the cytoplasmic domain(s) (CD); (c) isolated nucleotides that encode mutants, engineered or naturally occurring, of the described NGPCR in
30 which all or a part of at least one of the domains is deleted or altered, and the polypeptide products specified by such nucleotide sequences, including but not limited to soluble

receptors in which all or a portion of a TM is deleted, and nonfunctional receptors in which all or a portion of a CD is deleted; (d) nucleotides that encode fusion proteins containing the coding region from NGPCR, or one of its domains (e.g., an
5 extracellular domain) fused to another peptide or polypeptide.

The invention also encompasses agonists and antagonists of the NGPCR, including small molecules, large molecules, mutant NGPCR proteins, or portions thereof that compete with the native NGPCR, and antibodies, as well as nucleotide sequences that can
10 be used to inhibit the expression of the described NGPCR (e.g., antisense and ribozyme molecules, and gene or regulatory sequence replacement constructs) or to enhance the expression of the described NGPCR sequence (e.g., expression constructs that place the described sequence under the control of a strong promoter
15 system), and transgenic animals that express a NGPCR transgene or "knockouts" (which can be conditional) that do not express functional NGPCR (see, for example, PCT Applic. No. PCT/US98/03243, filed February 20, 1998, herein incorporated by reference). In addition to knockouts other aspects of the
20 present invention include animals having genetically engineered mutations in the described sequence that modify (i.e., point mutations, over-expression mutations, etc.) the activity or expression of the described NGPCR.

Further, the present invention also relates to methods of
25 using the described NGPCR coding region and/or NGPCR gene product(s) to identify compounds that modulate, i.e., act as agonists or antagonists, of NGPCR gene expression and/or NGPCR gene product activity. Such compounds can be used as therapeutic agents for the treatment of various symptomatic representations
30 of biological disorders or imbalances.

4. DESCRIPTION OF THE SEQUENCE LISTING AND FIGURES

The Sequence Listing provides the nucleotide sequence of the described NGPCR, and the amino acid sequence encoded thereby.

SEQ ID NO:3 shows the full length open reading frame and flanking
5 5' and 3' sequences.

5. DETAILED DESCRIPTION OF THE INVENTION

The human NGPCR described for the first time herein, is a novel receptor protein that is expressed in human cells. The NGPCR is a transmembrane protein that falls within the 7TM family
10 of receptors. As with other GPCRs, the NGPCR modulates signal transduction after the appropriate ligand has bound to the receptor. Interfering with the binding of the natural ligand, or neutralizing or removing the ligand, or interfering with its binding to NGPCR will effect NGPCR mediated signal transduction.

15 Because of their biological significance, 7TM, and particularly GPCR, proteins have been subjected to intense scientific/commercial scrutiny (see, for example, U.S. Applic. Ser. Nos. 08/820,521, filed March 19, 1997, and 08/833,226, filed April 17, 1997, both of which are herein incorporated by
20 reference in their entirety, which describe a variety of applications and uses for GPCR sequences such as the presently described NGPCR).

The invention encompasses the use of the described NGPCR nucleotides, NGPCR protein and peptides, as well as antibodies,
25 preferably humanized monoclonal antibodies, or binding fragments, domains, or fusion proteins thereof, to the NGPCR (which can, for example, act as NGPCR agonists or antagonists), antagonists that inhibit receptor activity or expression, or agonists that activate receptor activity or increase its expression in the
30 diagnosis and treatment of disease.

In particular, the invention described in the subsections below encompasses NGPCR polypeptides or peptides corresponding to

functional domains of the NGPCR (e.g., ECD, TM or CD), mutated, truncated or deleted forms of the NGPCR (e.g., modified versions missing one or more functional domains or portions thereof, such as, Δ ECD, Δ TM and/or Δ CD), NGPCR fusion proteins (e.g., NGPCR or
5 a functional domain of NGPCR, such as an ECD, fused to an unrelated protein or peptide such as an immunoglobulin constant region, i.e., IgFc), nucleotide sequences encoding such products, and host cell expression systems that can produce such NGPCR products.

10 The invention also encompasses antibodies and anti-idiotypic antibodies (including Fab fragments), antagonists and agonists of the NGPCR, as well as compounds or nucleotide constructs that inhibit the expression of NGPCR gene (transcription factor inhibitors, antisense and ribozyme molecules, or gene or
15 regulatory sequence replacement constructs), or promote expression of NGPCR (e.g., expression constructs in which NGPCR coding sequences are operatively associated with expression control elements such as promoters, promoter/enhancers, etc.). The invention also relates to host cells and animals genetically
20 engineered to express the human NGPCR (or mutant variants thereof) or to inhibit or "knockout" expression of an animal's endogenous NGPCR gene.

The NGPCR protein, or peptides therefrom, NGPCR fusion proteins, NGPCR nucleotide sequences, antibodies, antagonists and
25 agonists can be useful for the detection of mutant NGPCRs or inappropriately expressed variants of the NGPCR for the diagnosis of disease. The NGPCR protein, or peptides therefrom, NGPCR fusion proteins, NGPCR nucleotide sequences, host cell expression systems, antibodies, antagonists, agonists and genetically
30 engineered cells and animals can be used for screening for drugs (or high throughput screening of combinatorial libraries) effective in the treatment of the symptomatic or phenotypic manifestations of perturbing the normal function of NGPCR in the

body. The use of engineered host cells and/or animals may offer an advantage in that such systems allow not only for the identification of compounds that bind to an ECD of NGPCR, but can also identify compounds that affect the signal transduced by an
5 activated NGPCR.

Finally, the NGPCR protein products (especially soluble derivatives such as peptides corresponding to the NGPCR ECD, or truncated polypeptides lacking one or more TM domains) and fusion protein products (especially NGPCR-Ig fusion proteins, *i.e.*,
10 fusions of NGPCR, or a domain of NGPCR, *e.g.*, ECD, Δ TM to an IgFc), antibodies and anti-idiotypic antibodies (including Fab fragments), antagonists or agonists (including compounds that modulate signal transduction which may act on downstream targets in a NGPCR-mediated signal transduction pathway) can be used for
15 therapy of such diseases. For example, the administration of an effective amount of soluble NGPCR ECD, Δ TM, or an ECD-IgFc fusion protein or an anti-idiotypic antibody (or its Fab) that mimics the NGPCR ECD would "mop up" or "neutralize" the endogenous NGPCR ligand, and prevent or reduce binding and receptor activation.
20 Nucleotide constructs encoding such NGPCR products can be used to genetically engineer host cells to express such products *in vivo*; these genetically engineered cells function as "bioreactors" in the body delivering a continuous supply of NGPCR, a NGPCR peptide, soluble ECD or Δ TM or a NGPCR fusion protein that will
25 "mop up" or neutralize a NGPCR ligand. Nucleotide constructs encoding functional NGPCR, mutant NGPCR variants, as well as antisense and ribozyme molecules can be used in "gene therapy" approaches for the modulation of NGPCR expression. Thus, the invention also encompasses pharmaceutical formulations and
30 methods for treating biological disorders.

Various aspects of the invention are described in greater detail in the subsections below.

5.1 THE NGPCR GENE

The cDNA sequence (SEQ ID NO: 1) and deduced amino acid sequences (SEQ ID NO: 2) of human NGPCR are presented in the Sequence Listing. The described sequences were generated using
5 sequences obtained from gene trapped human cells and a human placenta and thyroid cDNA libraries (Edge Biosystems, Gaithersburg, MD). Expression studies have indicated that the described NGPCR can be expressed in human brain, cerebellum, spinal cord, thymus, spleen, bone marrow, liver, placenta,
10 prostate, thyroid, testis, adrenal gland, stomach, small intestine, colon, esophagus, bladder, rectum, pericardium, fetal lung, and gene trapped human cells. In view of the tissues that may express the NGPCR, the described human NGPCR may be an important target for therapeutic treatment of, *inter alia*,
15 diabetes, abnormal body weight or obesity, atherosclerosis, mental disorders, disorders of the immune system, heart disease, abnormal blood pressure, cancer, and any associated symptoms.

The NGPCR nucleotides of the present invention includes:
(a) the human DNA sequence presented in the Sequence Listing and
20 additionally contemplate any nucleotide sequences encoding a contiguous and functional NGPCR open reading frame (ORF) that hybridizes to a complement of the DNA sequence presented in the Sequence Listing under highly stringent conditions, *e.g.*, hybridization to filter-bound DNA in 0.5 M NaHPO₄, 7% sodium
25 dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1xSSC/0.1% SDS at 68°C (Ausubel F.M. *et al.*, eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York, at p. 2.10.3) and encodes a functionally equivalent gene product.
30 Additionally contemplated are any nucleotide sequences that hybridize to the complement of the DNA sequence that encodes and expresses an amino acid sequence presented in the Sequence Listing under moderately stringent conditions, *e.g.*, washing in

0.2xSSC/0.1% SDS at 42°C (Ausubel *et al.*, 1989, *supra*), yet which still encode a functionally equivalent NGPCR gene product. Functional equivalents of NGPCR include naturally occurring NGPCRs present in other species, and mutant NGPCRs whether
5 naturally occurring or engineered (by site directed mutagenesis, gene shuffling, directed evolution as described in, for example, U.S. Patent No. 5,837,458, herein incorporated by reference). The invention also includes degenerate variants of the disclosed sequence.

10 Additionally contemplated are polynucleotides encoding NGPCR ORFs, or their functional equivalents, encoded by polynucleotide sequences that are about 99, 95, 90, or about 85 percent similar or identical to corresponding regions of the nucleotide sequences of the Sequence Listing (as measured by BLAST sequence comparison
15 analysis using, for example, the GCG sequence analysis package (Madison, Wisconsin) using standard default settings).

 The invention also includes nucleic acid molecules, preferably DNA molecules, that hybridize to, and are therefore the complements of, the described NGPCR nucleotide sequence. Such
20 hybridization conditions may be highly stringent or less highly stringent, as described above. In instances wherein the nucleic acid molecules are deoxyoligonucleotides ("DNA oligos"), such molecules are particularly about 16 to about 100 bases long, about 20 to about 80, or about 34 to about 45 bases long, or any
25 variation or combination of sizes represented therein and which incorporate a contiguous region of sequence first disclosed in the Sequence Listing. The described oligonucleotides, can be used in conjunction with the polymerase chain reaction (PCR) to screen libraries, isolate clones, and prepare cloning and
30 sequencing templates, etc. Alternatively, such NGPCR oligonucleotides can be used as hybridization probes for screening libraries, and assessing gene expression patterns (particularly using a micro array or high-throughput "chip"

format). Additionally, a series of the described NGPCR oligonucleotide sequences, or the complements thereof, can be used to represent all or a portion of the described NGPCR sequences. An oligonucleotide or polynucleotide sequence first disclosed in at least a portion of one or more of the sequences of SEQ ID NOS: 1-3 can be used as a hybridization probe in conjunction with a solid support matrix/substrate (resins, beads, membranes, plastics, polymers, metal or metallized substrates, crystalline or polycrystalline substrates, etc.). Of particular note are spatially addressable arrays (i.e., gene chips, microtiter plates, etc.) of oligonucleotides and polynucleotides, or corresponding oligopeptides and polypeptides, wherein at least one of the biopolymers present on the spatially addressable array comprises an oligonucleotide or polynucleotide sequence first disclosed in at least one of the sequences of SEQ ID NOS: 1-3, or an amino acid sequence encoded thereby. Methods for attaching biopolymers to, or synthesizing biopolymers on, solid support matrices, and conducting binding studies thereon are disclosed in, *inter alia*, U.S. Patent Nos. 5,700,637, 5,556,752, 5,744,305, 4,631,211, 5,445,934, 5,252,743, 4,713,326, 5,424,186, and 4,689,405 the disclosures of which are herein incorporated by reference in their entirety.

Addressable arrays comprising sequences first disclosed in SEQ ID NOS:1-3 can be used to identify and characterize the temporal and tissue specific expression of a gene. These addressable arrays incorporate oligonucleotide sequences of sufficient length to confer the required specificity, yet be within the limitations of the production technology. The length of these probes is within a range of between about 8 to about 2000 nucleotides. Preferably the probes consist of 60 nucleotides and more preferably 25 nucleotides from the sequences first disclosed in SEQ ID NOS:1-3.

For example, a series of the described oligonucleotide

sequences, or the complements thereof, can be used in chip format to represent all or a portion of the described sequences. The oligonucleotides, typically between about 16 to about 40 (or any whole number within the stated range) nucleotides in length can partially overlap each other and/or the sequence may be represented using oligonucleotides that do not overlap. Accordingly, the described polynucleotide sequences shall typically comprise at least about two or three distinct oligonucleotide sequences of at least about 8 nucleotides in length that are each first disclosed in the described Sequence Listing. Such oligonucleotide sequences can begin at any nucleotide present within a sequence in the Sequence Listing and proceed in either a sense (5'-to-3') orientation vis-a-vis the described sequence or in an antisense orientation.

Microarray-based analysis allows the discovery of broad patterns of genetic activity, providing new understanding of gene functions and generating novel and unexpected insight into transcriptional processes and biological mechanisms. The use of addressable arrays comprising sequences first disclosed in SEQ ID NOS:1-3 provides detailed information about transcriptional changes involved in a specific pathway, potentially leading to the identification of novel components or gene functions that manifest themselves as novel phenotypes.

Probes consisting of sequences first disclosed in SEQ ID NOS:1-3 can also be used in the identification, selection and validation of novel molecular targets for drug discovery. The use of these unique sequences permits the direct confirmation of drug targets and recognition of drug dependent changes in gene expression that are modulated through pathways distinct from the drugs intended target. These unique sequences therefore also have utility in defining and monitoring both drug action and toxicity.

As an example of utility, the sequences first disclosed in

SEQ ID NOS:1-3 can be utilized in microarrays or other assay formats, to screen collections of genetic material from patients who have a particular medical condition. These investigations can also be carried out using the sequences first disclosed in
5 SEQ ID NOS:1-3 *in silico* and by comparing previously collected genetic databases and the disclosed sequences using computer software known to those in the art.

Thus the sequences first disclosed in SEQ ID NOS:1-3 can be used to identify mutations associated with a particular disease
10 and also as a diagnostic or prognostic assay.

Although the presently described sequences have been specifically described using nucleotide sequence, it should be appreciated that each of the sequences can uniquely be described using any of a wide variety of additional structural attributes,
15 or combinations thereof. For example, a given sequence can be described by the net composition of the nucleotides present within a given region of the sequence in conjunction with the presence of one or more specific oligonucleotide sequence(s) first disclosed in the SEQ ID NOS: 1-3. Alternatively, a
20 restriction map specifying the relative positions of restriction endonuclease digestion sites, or various palindromic or other specific oligonucleotide sequences can be used to structurally describe a given sequence. Such restriction maps, which are typically generated by widely available computer programs (e.g.,
25 the University of Wisconsin GCG sequence analysis package, SEQUENCHER 3.0, Gene Codes Corp., Ann Arbor, MI, etc.), can optionally be used in conjunction with one or more discrete nucleotide sequence(s) present in the sequence that can be described by the relative position of the sequence relative to one
30 or more additional sequence(s) or one or more restriction sites present in the disclosed sequence.

For oligonucleotide probes, highly stringent conditions may refer, e.g., to washing in 6xSSC/0.05% sodium pyrophosphate at

37°C (for 14-base oligos), 48°C (for 17-base oligos), 55°C (for 20-base oligos), and 60°C (for 23-base oligos). These nucleic acid molecules may encode or act as NGPCR antisense molecules, useful, for example, in NGPCR gene regulation (for and/or as antisense primers in amplification reactions of NGPCR gene nucleic acid sequences). With respect to NGPCR gene regulation, such techniques can be used to regulate biological functions. Further, such sequences may be used as part of ribozyme and/or triple helix sequences, also useful for NGPCR gene regulation.

10 Additionally, the antisense oligonucleotides may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil,

15 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine,

20 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine,

25 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

30 The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

 In yet another embodiment, the antisense oligonucleotide

comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or
5 analog thereof.

In yet another embodiment, the antisense oligonucleotide is an α -anomeric oligonucleotide. An α -anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to
10 each other (Gautier et al., 1987, Nucl. Acids Res. 15:6625-6641). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al., 1987, Nucl. Acids Res. 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue et al., 1987, FEBS Lett. 215:327-330).

Oligonucleotides of the invention may be synthesized by
15 standard methods known in the art, e.g. by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein et al. (1988, Nucl. Acids Res. 16:3209),
20 methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., 1988, Proc. Natl. Acad. Sci. U.S.A. 85:7448-7451), etc.

Low stringency conditions are well known to those of skill in the art, and will vary predictably depending on the specific
25 organisms from which the library and the labeled sequences are derived. For guidance regarding such conditions see, for example, Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual (and periodic updates thereof), Cold Springs Harbor Press, N.Y.; and Ausubel et al., 1989, Current Protocols in Molecular
30 Biology, Green Publishing Associates and Wiley Interscience, N.Y.

Alternatively, suitably labeled NGPCR nucleotide probes may be used to screen a human genomic library using appropriately

stringent conditions or by PCR. The identification and characterization of human genomic clones is helpful for identifying polymorphisms, determining the genomic structure of a given locus/allele, and designing diagnostic tests. For example, sequences derived from regions adjacent to the intron/exon boundaries of the human gene can be used to design primers for use in amplification assays to detect mutations within the exons, introns, splice sites (e.g., splice acceptor and/or donor sites), etc., that can be used in diagnostics and pharmacogenomics.

10 Further, a NGPCR gene homolog can be isolated from nucleic acid from the organism of interest by performing PCR using two degenerate oligonucleotide primer pools designed on the basis of amino acid sequences within the NGPCR sequence product disclosed herein. The template for the reaction may be total RNA, mRNA, and/or cDNA obtained by reverse transcription of mRNA prepared from, for example, human or non-human cell lines or tissue, such as choroid plexus, known or suspected to express a NGPCR gene allele.

The PCR product may be subcloned and sequenced to ensure that the amplified sequences represent the sequence of the desired NGPCR gene. The PCR fragment may then be used to isolate a full length cDNA clone by a variety of methods. For example, the amplified fragment may be labeled and used to screen a cDNA library, such as a bacteriophage cDNA library. Alternatively, the labeled fragment may be used to isolate genomic clones via the screening of a genomic library.

PCR technology may also be utilized to isolate full length cDNA sequences. For example, RNA may be isolated, following standard procedures, from an appropriate cellular or tissue source (i.e., one known, or suspected, to express the NGPCR gene, such as, for example, adipose tissue/fat cells). A reverse transcription (RT) reaction may be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the

amplified fragment for the priming of first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed" using a standard terminal transferase reaction, the hybrid may be digested with RNase H, and second strand synthesis may then be primed with a complementary primer. Thus, cDNA sequences upstream of the amplified fragment may easily be isolated. For a review of cloning strategies which may be used, see e.g., Sambrook *et al.*, 1989, *supra*.

A cDNA of a mutant NGPCR gene may be isolated, for example, by using PCR. In this case, the first cDNA strand may be synthesized by hybridizing an oligo-dT oligonucleotide to mRNA isolated from tissue known or suspected to be expressed in an individual putatively carrying a mutant NGPCR allele, and by extending the new strand with reverse transcriptase. The second strand of the cDNA is then synthesized using an oligonucleotide that hybridizes specifically to the 5' end of the normal gene. Using these two primers, the product is then amplified via PCR, optionally cloned into a suitable vector, and subjected to DNA sequence analysis through methods well known to those of skill in the art. By comparing the DNA sequence of the mutant NGPCR allele to that of the normal NGPCR allele, the mutation(s) responsible for the loss or alteration of function of the mutant NGPCR gene product can be ascertained.

Alternatively, a genomic library can be constructed using DNA obtained from an individual suspected of or known to carry the mutant NGPCR allele, or a cDNA library can be constructed using RNA from a tissue known, or suspected, to express the mutant NGPCR allele. A normal NGPCR gene, or any suitable fragment thereof, can then be labeled and used as a probe to identify the corresponding mutant NGPCR allele in such libraries. Clones containing the mutant NGPCR sequences may then be purified and subjected to sequence analysis according to methods well known to those of skill in the art.

Additionally, an expression library can be constructed utilizing cDNA synthesized from, for example, RNA isolated from a tissue known, or suspected, to express a mutant NGPCR allele in an individual suspected of or known to carry such a mutant allele. In this manner, gene products made by the putatively mutant tissue may be expressed and screened using standard antibody screening techniques in conjunction with antibodies raised against the normal NGPCR gene product, as described, below, in Section 5.3. (For screening techniques, see, for example, Harlow, E. and Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor.)

Additionally, screening can be accomplished by screening with labeled NGPCR fusion proteins, such as, for example, alkaline phosphatase-NGPCR or NGPCR-alkaline phosphatase fusion proteins. In cases where a NGPCR mutation results in an expressed gene product with altered function (e.g., as a result of a missense or a frameshift mutation), a polyclonal set of antibodies to NGPCR are likely to cross-react with the mutant NGPCR gene product. Library clones detected via their reaction with such labeled antibodies can be purified and subjected to sequence analysis according to methods well known to those of skill in the art.

The invention also encompasses nucleotide sequences that encode mutant NGPCRs, peptide fragments of the NGPCR, truncated NGPCRs, and NGPCR fusion proteins. These include, but are not limited to nucleotide sequences encoding mutant NGPCRs described in section 5.2 *infra*; polypeptides or peptides corresponding to one or more ECD, TM and/or CD domains of the NGPCR or portions of these domains; truncated NGPCRs in which one or two of the domains is deleted, e.g., a soluble NGPCR lacking the TM or both the TM and CD regions, or a truncated, nonfunctional NGPCR lacking all or a portion of a CD, for example, region.

Nucleotides encoding fusion proteins may include, but are not limited to, full length NGPCR sequence, truncated NGPCRs, or nucleotides encoding peptide fragments of NGPCR fused to an unrelated protein or peptide, such as for example, a

5 transmembrane sequence, which anchors the NGPCR ECD to the cell membrane; an Ig Fc domain which increases the stability and half life of the resulting fusion protein (e.g., NGPCR-Ig) in the bloodstream; or an enzyme, fluorescent protein, luminescent protein which can be used as a marker.

10 The invention also encompasses (a) DNA vectors that contain any of the NGPCR coding sequence and/or the complement thereof (i.e., antisense); (b) DNA expression vectors that contain any of the NGPCR coding sequence operatively associated with a regulatory element that directs the expression of the coding

15 sequence; and (c) genetically engineered host cells engineered to contain NGPCR coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences in the host cell. As used herein, regulatory elements include, but are not limited to, inducible and non-inducible

20 promoters, enhancers, operators and other elements known to those skilled in the art that drive and regulate expression. Such regulatory elements include but are not limited to the human cytomegalovirus (hCMV) immediate early gene, regulatable, viral elements (particularly retroviral LTR promoters), the early or

25 late promoters of SV40 adenovirus, the *lac* system, the *trp* system, the *TAC* system, the *TRC* system, the major operator and promoter regions of phage lambda, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase (PGK), the promoters of acid phosphatase, and the promoters of the yeast

30 α -mating factors.

5.2 NGPCR PROTEIN

NGPCR protein, peptide fragments therefrom, mutated, truncated or deleted forms of the NGPCR and/or NGPCR fusion proteins can be prepared for a variety of uses, including but not
5 limited to the generation of antibodies, as reagents in diagnostic assays, the identification of other cellular gene products related to the NGPCR, as reagents in assays for screening for compounds that can be as pharmaceutical reagents useful in the therapeutic treatment of mental, biological, or
10 medical disorders and disease.

The Sequence Listing discloses the amino acid sequence encoded by the described NGPCR ORF. The described NGPCR has an initiator methionine in a DNA sequence context consistent with a translation initiation site, followed by a initiator codon.

15 The NGPCR sequence of the present invention includes the nucleotide and amino acid sequences presented in the Sequence Listing as well as analogues and derivatives thereof. Further, corresponding NGPCR homologues from other species are encompassed by the invention. In fact, any NGPCR protein encoded by the
20 NGPCR nucleotides described above, is within the scope of the invention, as are any novel polynucleotide sequences encoding all or any novel portion of an amino acid sequence presented in the Sequence Listing. The degenerate nature of the genetic code is well known, and, accordingly, each amino acid presented in the
25 Sequence Listing, is generically representative of the well known nucleic acid "triplet" codon, or in many cases codons, that can encode the amino acid. As such, as contemplated herein, the amino acid sequence presented in the Sequence Listing, when taken together with the genetic code (see, for example, Table 4-1 at
30 page 109 of "Molecular Cell Biology", 1986, J. Darnell et al. eds., Scientific American Books, New York, NY, herein incorporated by reference) are generically representative of all the various permutations and combinations of nucleic acid

sequences that can encode such amino acid sequence (as well as such variants as biased by human codon usage frequency tables).

The invention also encompasses proteins that are functionally equivalent to the NGPCR encoded by the nucleotide sequence described in Section 5.1, as judged by any of a number of criteria, including but not limited to the ability to bind a ligand of the NGPCR, the ability to effect an identical or complementary signal transduction pathway, a change in cellular metabolism (e.g., ion flux, tyrosine phosphorylation, etc.), or effect the same change in phenotype when the NGPCR equivalent is present in an appropriate cell type (such as the amelioration, prevention or delay of a biochemical, biophysical, or overt phenotype). Such functionally equivalent NGPCR proteins include but are not limited to additions or substitutions of amino acid residues within the amino acid sequence encoded by the NGPCR nucleotide sequence described above, in Section 5.1, but which result in a silent change, thus producing a functionally equivalent gene product. Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

While random mutations can be made to NGPCR DNA (using random mutagenesis techniques well known to those skilled in the art) and the resulting mutant NGPCRs tested for activity, site-directed mutations of the NGPCR coding sequence can be engineered (using site-directed mutagenesis techniques well known to those skilled in the art) to generate mutant NGPCRs with increased

function, e.g., higher binding affinity for the target ligand, and/or greater signaling capacity; or decreased function, and/or decreased signal transduction capacity. One starting point for such analysis is by aligning the disclosed human sequences with
5 corresponding gene/protein sequences from, for example, other mammals in order to identify amino acid sequence motifs that are conserved between different species. Non-conservative changes can be engineered at variable positions to alter function, signal transduction capability, or both. Alternatively, where
10 alteration of function is desired, deletion or non-conservative alterations of the conserved regions (i.e., identical amino acids) can be engineered. For example, deletion or non-conservative alterations (substitutions or insertions) of the various conserved transmembrane domains.

15 Other mutations to the NGPCR coding sequence can be made to generate NGPCRs that are better suited for expression, scale up, etc. in the host cells chosen. For example, cysteine residues can be deleted or substituted with another amino acid in order to eliminate disulfide bridges; N-linked glycosylation sites can be
20 altered or eliminated to achieve, for example, expression of a homogeneous product that is more easily recovered and purified from yeast hosts which are known to hyperglycosylate N-linked sites. To this end, a variety of amino acid substitutions at one or both of the first or third amino acid positions of any one or
25 more of the glycosylation recognition sequences which occur in an ECD (N-X-S or N-X-T), and/or an amino acid deletion at the second position of any one or more such recognition sequences in an ECD will prevent glycosylation of the NGPCR at the modified tripeptide sequence. (See, e.g., Miyajima *et al.*, 1986, EMBO J.
30 5(6):1193-1197).

Peptides corresponding to one or more domains of the NGPCR (e.g., ECD, TM, CD, etc.), truncated or deleted NGPCRs (e.g., NGPCR in which a ECD, TM and/or CD is deleted) as well as fusion

proteins in which a full length NGPCR, a NGPCR peptide, or truncated NGPCR is fused to an unrelated protein, are also within the scope of the invention, and can be designed on the basis of the presently disclosed NGPCR nucleotide and NGPCR amino acid sequences. Such fusion proteins include but are not limited to IgFc fusions which stabilize the NGPCR protein or peptide and prolong half-life *in vivo*; or fusions to any amino acid sequence that allows the fusion protein to be anchored to the cell membrane, allowing an ECD to be exhibited on the cell surface; or fusions to an enzyme, fluorescent protein, or luminescent protein which provide a marker function.

While the NGPCR polypeptide, and corresponding peptides, can be chemically synthesized (e.g., see Creighton, 1983, *Proteins: Structures and Molecular Principles*, W.H. Freeman & Co., N.Y.), large polypeptides derived from NGPCR and especially full-length NGPCR product can be advantageously produced by recombinant DNA technology using techniques well known in the art for expressing nucleic acid containing NGPCR sequences and/or coding sequences. Such methods can be used to construct expression vectors containing the NGPCR nucleotide sequence(s) described in Section 5.1 and appropriate transcriptional and translational control signals. These methods include, for example, *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. See, for example, the techniques described in Sambrook et al., 1989, *supra*, and Ausubel et al., 1989, *supra*. Alternatively, RNA corresponding to all or a portion of a transcript encoded by the NGPCR nucleotide sequence may be chemically synthesized using, for example, synthesizers. See, for example, the techniques described in "Oligonucleotide Synthesis", 1984, Gait, M.J. ed., IRL Press, Oxford, which is incorporated by reference herein in its entirety.

A variety of host-expression vector systems may be utilized to express the NGPCR nucleotide sequence of the invention. Where

the NGPCR peptide or polypeptide is a soluble derivative (e.g., NGPCR peptides corresponding to an ECD; truncated or deleted NGPCR in which a TM and/or CD are deleted) the peptide or polypeptide can be recovered from the culture, i.e., from the host cell in cases where the NGPCR peptide or polypeptide is not secreted, and from the culture media in cases where the NGPCR peptide or polypeptide is secreted by the cells. However, such expression systems also encompass engineered host cells that express NGPCR, or functional equivalent, *in situ*, i.e., anchored in the cell membrane. Purification or enrichment of NGPCR from such expression systems can be accomplished using appropriate detergents and lipid micelles and methods well known to those skilled in the art. However, such engineered host cells themselves may be used in situations where it is important not only to retain the structural and functional characteristics of the NGPCR, but to assess biological activity, e.g., in drug screening assays.

The expression systems that may be used for purposes of the invention include but are not limited to microorganisms such as bacteria (e.g., *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing the NGPCR nucleotide sequence; yeast (e.g., *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing the NGPCR nucleotide sequence; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing the NGPCR sequence; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing the NGPCR nucleotide sequence; or mammalian cell systems (e.g., COS, CHO, BHK, 293, 3T3, etc.) harboring recombinant expression constructs containing promoters

derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter, etc.).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the NGPCR gene product being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions of NGPCR protein, or for raising antibodies to the NGPCR protein, or corresponding peptide, for example, vectors that direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther et al., 1983, EMBO J. 2:1791), in which the NGPCR coding sequence may be ligated individually into the vector in frame with the *lacZ* coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, 1985, Nucleic Acids Res. 13:3101-3109; Van Heeke & Schuster, 1989, J. Biol. Chem. 264:5503-5509); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The PGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

In an insect system, *Autographa californica* nuclear polyhidrosis virus (AcNPV) is used as a vector to express foreign sequences. The virus grows in *Spodoptera frugiperda* cells. The NGPCR coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the

polyhedrin promoter). Successful insertion of NGPCR gene coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (*i.e.*, virus lacking the proteinaceous coat coded for by the polyhedrin gene). These
5 recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed (*e.g.*, see Smith *et al.*, 1983, *J. Virol.* 46: 584; Smith, U.S. Patent No. 4,215,051).

In mammalian host cells, a number of viral-based expression
10 systems may be utilized. In cases where an adenovirus is used as an expression vector, the NGPCR nucleotide sequence may be ligated to an adenovirus transcription/translation control complex, *e.g.*, the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome
15 by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (*e.g.*, region E1 or E3) will result in a recombinant virus that is viable and capable of expressing the NGPCR gene product in infected hosts (*e.g.*, See Logan & Shenk, 1984, *Proc. Natl. Acad. Sci. USA* 81:3655-3659).
20 Specific initiation signals may also be required for efficient translation of inserted NGPCR nucleotide sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire NGPCR gene or cDNA, including its own initiation codon and adjacent sequences, is inserted into the
25 appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the NGPCR coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation
30 codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and

synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (See Bittner et al., 1987, Methods in Enzymol. 153:516-544).

5 In addition, a host cell strain may be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function
10 of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To
15 this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include, but are not limited to, CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, and in
20 particular, choroid plexus cell lines.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express a NGPCR sequence may be engineered. Rather than using expression vectors which contain viral origins
25 of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the
introduction of the foreign DNA, engineered cells may be allowed
30 to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to

form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which express the NGPCR gene product. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the NGPCR gene product.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, et al., 1977, Cell 11:223), hypoxanthine-guanine phosphoribosyltransferase (Szybalski & Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48:2026), and adenine phosphoribosyltransferase (Lowy, et al., 1980, Cell 22:817) genes can be employed in tk⁻, hgp⁻ or ap⁻ cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for the following genes: dhfr, which confers resistance to methotrexate (Wigler, et al., 1980, Natl. Acad. Sci. USA 77:3567; O'Hare, et al., 1981, Proc. Natl. Acad. Sci. USA 78:1527); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, 1981, Proc. Natl. Acad. Sci. USA 78:2072); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, et al., 1981, J. Mol. Biol. 150:1); and hyg^r, which confers resistance to hygromycin (Santerre, et al., 1984, Gene 30:147).

Alternatively, any fusion protein may be readily purified by utilizing an antibody specific for the fusion protein being expressed. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht, et al., 1991, Proc. Natl. Acad. Sci. USA 88: 8972-8976). In this system, the sequence of interest is subcloned into a vaccinia recombination plasmid such that the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six

histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni²⁺-nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

5 NGPCR gene product(s) can also be expressed in transgenic animals. Animals of any species, including, but not limited to, worms, mice, rats, rabbits, guinea pigs, pigs, micro-pigs, birds, goats, and non-human primates, e.g., baboons, monkeys, and chimpanzees may be used to generate NGPCR transgenic animals.

10 Any technique known in the art may be used to introduce a NGPCR transgene into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to pronuclear microinjection (Hoppe, P.C. and Wagner, T.E., 1989, U.S. Pat. No. 4,873,191); retrovirus mediated gene transfer into
15 germline cells (Van der Putten *et al.*, 1985, Proc. Natl. Acad. Sci., USA 82:6148-6152); gene targeting in embryonic stem cells (Thompson *et al.*, 1989, Cell 56:313-321); electroporation of embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814); and sperm-mediated gene transfer (Lavitrano *et al.*, 1989, Cell 57:717-723);
20 etc. For a review of such techniques, see Gordon, 1989, Transgenic Animals, Intl. Rev. Cytol. 115:171-229, which is incorporated by reference herein in its entirety.

 The present invention provides for transgenic animals that carry the NGPCR transgene in all their cells, as well as animals
25 which carry the transgene in some, but not all their cells, i.e., mosaic animals or somatic cell transgenic animals. The transgene may be integrated as a single transgene or in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular
30 cell type by following, for example, the teaching of Lasko *et al.*, 1992, Proc. Natl. Acad. Sci. USA 89:6232-6236. The regulatory sequences required for such a cell-type specific

activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

When it is desired that a NGPCR transgene be integrated into the chromosomal site of the endogenous NGPCR gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous NGPCR gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous NGPCR gene (*i.e.*, "knockout" animals).

The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous NGPCR gene in only that cell type, by following, for example, the teaching of Gu *et al.*, 1994, *Science*, 265:103-106. The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant NGPCR gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to assay whether integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include but are not limited to Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and RT-PCR. Samples of NGPCR gene-expressing tissue, may also be evaluated immunocytochemically using antibodies specific for the NGPCR transgene product.

5.3 ANTIBODIES TO NGPCR PROTEINS

Antibodies that specifically recognize one or more epitopes of NGPCR, or epitopes of conserved variants of NGPCR, or peptide fragments of NGPCR are also encompassed by the invention. Such
5 antibodies include but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')₂ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the
10 above.

The antibodies of the invention may be used, for example, in the detection of NGPCR in a biological sample and may, therefore, be utilized as part of a diagnostic or prognostic technique whereby patients may be tested for abnormal amounts of NGPCR.
15 Such antibodies may also be utilized in conjunction with, for example, compound screening schemes, as described, below, in Section 5.5, for the evaluation of the effect of test compounds on expression and/or activity of a NGPCR gene product. Additionally, such antibodies can be used in conjunction with
20 gene therapy to, for example, evaluate the normal and/or engineered NGPCR-expressing cells prior to their introduction into the patient. Such antibodies may additionally be used to stimulate or inhibit NGPCR activity. Thus, such antibodies may, therefore, be utilized as part of treatment methods for NGPCR-
25 related biological disorders.

For the production of antibodies, various host animals may be immunized by injection with the NGPCR, a NGPCR peptide (e.g., one corresponding to a functional domain of the receptor, such as an ECD, TM or CD), truncated NGPCR polypeptides (NGPCR in which
30 one or more domains, e.g., a TM or CD, has been deleted), functional equivalents of the NGPCR or mutants of the NGPCR. Such host animals may include, but are not limited to, rabbits, mice, and rats, to name but a few. Various adjuvants may be used

to increase the immunological response, depending on the host species, including but not limited to Freund's adjuvant (complete and incomplete), mineral salts such as aluminum hydroxide or aluminum phosphate, surface active substances such as

5 lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*.

Alternatively, the immune response could be enhanced by combination and or coupling with molecules such as keyhole limpet
10 hemocyanin, tetanus toxoid, diptheria toxoid, ovalbumin, cholera toxin or fragments thereof. Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of the immunized animals.

Monoclonal antibodies, which are homogeneous populations of
15 antibodies to a particular antigen, may be obtained by any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique of Kohler and Milstein, (1975, Nature 256:495-497; and U.S. Patent No. 4,376,110), the
20 human B-cell hybridoma technique (Kosbor et al., 1983, Immunology Today 4:72; Cole et al., 1983, Proc. Natl. Acad. Sci. USA 80:2026-2030), and the EBV-hybridoma technique (Cole et al., 1985, Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin
25 class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated in vitro or in vivo. Production of high titers of mAbs in vivo makes this the presently preferred method of production. Also favored is the production of fully humanized
30 monoclonal antibodies as described in US Patent No. 6,150,584 and respective disclosures which are herein incorporated by reference in their entirety.

In addition, techniques developed for the production of

"chimeric antibodies" (Morrison et al., 1984, Proc. Natl. Acad. Sci., 81:6851-6855; Neuberger et al., 1984, Nature, 312:604-608; Takeda et al., 1985, Nature, 314:452-454) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region. Such technologies are described in U.S. Patents Nos. 6,075,181 and 5,877,397 and their respective disclosures which are herein incorporated by reference in their entirety.

Alternatively, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778; Bird, 1988, Science 242:423-426; Huston et al., 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; and Ward et al., 1989, Nature 334:544-546) can be adapted to produce single chain antibodies against NGPCR gene products. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, such fragments include but are not limited to: the F(ab')₂ fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed (Huse et al., 1989, Science, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

Antibodies to NGPCR can be used as substitutes for NGPCR ligands, to inhibit the binding of other ligands, or in turn be

utilized to generate anti-idiotypic antibodies that "mimic" NGPCR, using techniques well known to those skilled in the art. (See, e.g., Greenspan & Bona, 1993, FASEB J 7(5):437-444; and Nissinoff, 1991, J. Immunol. 147(8):2429-2438). For example

5 antibodies which bind to a NGPCR ECD and competitively inhibit the binding of a ligand of NGPCR can be used to generate anti-idiotypes that "mimic" a NGPCR ECD and, therefore, bind and neutralize a ligand. Such neutralizing anti-idiotypes or Fab fragments of such anti-idiotypes can be used in therapeutic

10 regimens involving the NGPCR signaling pathway.

5.4 DIAGNOSIS OF ABNORMALITIES RELATED TO NGPCR

A variety of methods can be employed for the diagnostic and prognostic evaluation of disorders related to NGPCR function, and

15 for the identification of subjects having a predisposition to such disorders.

Such methods may, for example, utilize reagents such as the NGPCR nucleotide sequence described in Section 5.1, or a portion thereof, and NGPCR antibodies as described in Section 5.3.

20 Specifically, such reagents may be used, for example, for:

(1) the detection of the presence of NGPCR gene mutations, or the detection of either over- or under-expression of NGPCR mRNA relative to a given phenotype; (2) the detection of either an over- or an under-abundance of NGPCR gene product relative to a

25 given phenotype; and (3) the detection of perturbations or abnormalities in the signal transduction pathway mediated by NGPCR.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one

30 specific NGPCR nucleotide sequence or NGPCR antibody reagent described herein, which may be conveniently used, e.g., in clinical settings, to diagnose patients exhibiting biological abnormalities.

For the detection of NGPCR mutations, any nucleated cell can be used as a starting source for genomic nucleic acid. For the detection of NGPCR gene expression or NGPCR gene products, any cell type or tissue in which the NGPCR gene is expressed, such as, for example, adipocytes, may be utilized.

Nucleic acid-based detection techniques are described, below, in Section 5.4.1. Peptide detection techniques are described, below, in Section 5.4.2.

5.4.1 DETECTION OF THE NGPCR GENE AND TRANSCRIPT

10 Mutations within the NGPCR gene can be detected by utilizing a number of techniques. Nucleic acid from any nucleated cell can be used as the starting point for such assay techniques, and may be isolated according to standard nucleic acid preparation procedures which are well known to those of skill in the art.

15 DNA may be used in hybridization or amplification assays of biological samples to detect abnormalities involving NGPCR gene structure, including point mutations, insertions, deletions and chromosomal rearrangements. Such assays may include, but are not limited to, Southern analyses, single stranded conformational polymorphism analyses (SSCP), and PCR analyses.

Such diagnostic methods for the detection of NGPCR gene-specific mutations can involve for example, contacting and incubating nucleic acids including recombinant DNA molecules, cloned genes or degenerate variants thereof, obtained from a sample, e.g., derived from a patient sample or other appropriate cellular source, with one or more labeled nucleic acid reagents including recombinant DNA molecules, cloned genes or degenerate variants thereof, as described in Section 5.1, under conditions favorable for the specific annealing of these reagents to their complementary sequences within the NGPCR gene. Preferably, the lengths of these nucleic acid reagents are at least 15 to 30 nucleotides. After incubation, all non-annealed nucleic acids

are removed from the nucleic acid:NGPCR molecule hybrid. The presence of nucleic acids which have hybridized, if any such molecules exist, is then detected. Using such a detection scheme, the nucleic acid from the cell type or tissue of interest
5 can be immobilized, for example, to a solid support such as a membrane, or a plastic surface such as that on a microtiter plate or polystyrene beads. In this case, after incubation, non-annealed, labeled nucleic acid reagents of the type described in Section 5.1 are easily removed. Detection of the remaining,
10 annealed, labeled NGPCR nucleic acid reagents is accomplished using standard techniques well-known to those in the art. The NGPCR gene sequence(s) to which the nucleic acid reagents have annealed can be compared to the annealing pattern expected from a normal NGPCR gene sequence in order to determine whether a NGPCR
15 gene mutation is present.

Alternative diagnostic methods for the detection of NGPCR gene specific nucleic acid molecules, in patient samples or other appropriate cell sources, may involve their amplification, e.g., by PCR (the experimental embodiment set forth in Mullis, K.B.,
20 1987, U.S. Patent No. 4,683,202), followed by the detection of the amplified molecules using techniques well known to those of skill in the art. The resulting amplified sequences can be compared to those which would be expected if the nucleic acid being amplified contained only normal copies of the NGPCR gene in
25 order to determine whether a NGPCR gene mutation exists.

Additionally, well-known genotyping techniques can be performed to identify individuals carrying NGPCR gene mutations. Such techniques include, for example, the use of restriction fragment length polymorphisms (RFLPs), which involve sequence
30 variations in one of the recognition sites for the specific restriction enzyme used.

Additionally, improved methods for analyzing DNA polymorphisms which can be utilized for the identification of

NGPCR gene mutations have been described which capitalize on the presence of variable numbers of short, tandemly repeated DNA sequences between the restriction enzyme sites. For example, Weber (U.S. Pat. No. 5,075,217, which is incorporated herein by reference in its entirety) describes a DNA marker based on length polymorphisms in blocks of (dC-dA)_n-(dG-dT)_n short tandem repeats. The average separation of (dC-dA)_n-(dG-dT)_n blocks is estimated to be 30,000-60,000 bp. Markers which are so closely spaced exhibit a high frequency co-inheritance, and are extremely useful in the identification of genetic mutations, such as, for example, mutations within the NGPCR gene, and the diagnosis of diseases and disorders related to such NGPCR mutations. Alternatively, single nucleotide polymorphisms (SNPs) or coding region SNPs (or CSNPs) can be identified for the disclosed NGPCR sequences.

Also, Caskey *et al.* (U.S. Pat. No. 5,364,759, which is incorporated herein by reference in its entirety) describe a DNA profiling assay for detecting short tri and tetra nucleotide repeat sequences. The process includes extracting the DNA of interest, such as the NGPCR gene, amplifying the extracted DNA, and labeling the repeat sequences to form a genotypic map of the individual's DNA.

The level of NGPCR gene expression can also be assayed by detecting and measuring NGPCR transcription. For example, RNA from a cell type or tissue known, or suspected to express the NGPCR gene, such as fat cells, may be isolated and tested utilizing hybridization or PCR techniques such as are described, above. The isolated cells can be derived from cell culture or from a patient. The analysis of cells taken from culture may be a necessary step in the assessment of cells to be used as part of a cell-based gene therapy technique or, alternatively, to test the effect of compounds on the expression of the NGPCR gene. Such analyses may reveal both quantitative and qualitative

aspects of the expression pattern of the NGPCR gene, including activation or inactivation of NGPCR gene expression.

In one embodiment of a NGPCR detection scheme, cDNAs are synthesized from the RNAs of interest (e.g., by reverse transcription of the RNA molecule into cDNA). A sequence within the cDNA is then used as the template for a nucleic acid amplification reaction, such as a PCR amplification reaction, or the like. The nucleic acid reagents used as synthesis initiation reagents (e.g., primers) in the reverse transcription and nucleic acid amplification steps of this method are chosen from among the NGPCR nucleic acid reagents described in Section 5.1. The preferred lengths of such nucleic acid reagents are at least 9-30 nucleotides. For detection of the amplified product, the nucleic acid amplification may be performed using radioactively or non-radioactively labeled nucleotides. Alternatively, enough amplified product may be made such that the product may be visualized by standard ethidium bromide staining, by utilizing any other suitable nucleic acid staining method, or by sequencing.

Additionally, it is possible to perform such NGPCR gene expression assays "in situ", i.e., directly upon tissue sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections, such that no nucleic acid purification is necessary. Nucleic acid reagents such as those described in Section 5.1 may be used as probes and/or primers for such in situ procedures (See, for example, Nuovo, G.J., 1992, "PCR In Situ Hybridization: Protocols And Applications", Raven Press, NY).

Alternatively, if a sufficient quantity of the appropriate cells can be obtained, standard Northern analysis can be performed to determine the level of mRNA expression of the NGPCR gene.

5.4.2 DETECTION OF NGPCR GENE PRODUCTS

Antibodies directed against wild type or mutant NGPCR gene products or conserved variants or peptide fragments thereof, which are discussed, above, in Section 5.3, may also be used as
5 diagnostics and prognostics, as described herein. Such diagnostic methods, may be used to detect abnormalities in the level of NGPCR gene expression, or abnormalities in the structure and/or temporal, tissue, cellular, or subcellular location of the NGPCR, and may be performed *in vivo* or *in vitro*, such as, for
10 example, on biopsy tissue.

For example, antibodies directed to epitopes of a NGPCR ECD can be used *in vivo* to detect the pattern and level of expression of NGPCR in the body. Such antibodies can be labeled, e.g., with a radio-opaque or other appropriate compound and injected into a
15 subject in order to visualize binding to the NGPCR expressed in the body using methods such as X-rays, CAT-scans, or MRI. Labeled antibody fragments, e.g., the Fab or single chain antibody comprising the smallest portion of the antigen binding region, are preferred for this purpose to promote crossing the
20 blood-brain barrier and permit labeling a NGPCR that may be expressed in the brain.

Additionally, any NGPCR fusion protein or NGPCR conjugated protein whose presence can be detected, can be administered. For example, NGPCR fusion or conjugated proteins labeled with a
25 radio-opaque or other appropriate compound can be administered and visualized *in vivo*, as discussed, above for labeled antibodies. Further such NGPCR fusion proteins as AP-NGPCR on NGPCR-Ap fusion proteins can be utilized for *in vitro* diagnostic procedures.

30 Alternatively, immunoassays or fusion protein detection assays, as described above, can be utilized on biopsy and autopsy samples *in vitro* to permit assessment of the expression pattern of the NGPCR. Such assays are not confined to the use of

antibodies that define a NGPCR ECD, but can include the use of antibodies directed to epitopes of any of the domains of NGPCR, e.g., a ECD, a TM and/or CD. The use of each or all of these labeled antibodies will yield useful information regarding translation and intracellular transport of the NGPCR to the cell surface, and can identify defects in processing.

The tissue or cell type to be analyzed will generally include those which are known, or suspected, to express the NGPCR gene, such as, for example, fat cells. The protein isolation methods employed herein may, for example, be such as those described in Harlow and Lane (Harlow, E. and Lane, D., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York), which is incorporated herein by reference in its entirety. The isolated cells can be derived from cell culture or from a patient. The analysis of cells taken from culture may be a necessary step in the assessment of cells that could be used as part of a cell-based gene therapy technique or, alternatively, to test the effects of various compounds on NGPCR gene expression.

For example, antibodies, or fragments of antibodies, such as those described, above, in Section 5.3, useful in the present invention may be used to quantitatively or qualitatively detect the presence of NGPCR gene products or conserved variants or peptide fragments thereof. This can be accomplished, for example, by immunofluorescence techniques employing a fluorescently labeled antibody (see below, this Section) coupled with light microscopic, flow cytometric, or fluorimetric detection. Such techniques are especially preferred if such NGPCR gene products are expressed on the cell surface.

The antibodies (or fragments thereof) or NGPCR fusion or conjugated proteins useful in the present invention may, additionally, be employed histologically, as in immunofluorescence, immunoelectron microscopy or non-immuno

assays, for *in situ* detection of NGPCR gene products or conserved variants or peptide fragments thereof, or for NGPCR binding (in the case of labeled NGPCR ligand fusion protein). *In situ* detection may be accomplished by removing a histological specimen
5 from a patient, and applying thereto a labeled antibody or fusion protein of the present invention. The antibody (or fragment) or fusion protein is preferably applied by overlaying the labeled antibody (or fragment) onto a biological sample. Through the use of such a procedure, it is possible to determine not only the
10 presence of the NGPCR gene product, or conserved variants or peptide fragments, or NGPCR binding, but also its distribution in the examined tissue. Using the present invention, those of ordinary skill will readily perceive that any of a wide variety of histological methods (such as staining procedures) can be
15 modified in order to achieve such *in situ* detection.

Immunoassays and non-immunoassays for NGPCR gene products or conserved variants or peptide fragments thereof will typically comprise incubating a sample, such as a biological fluid, a tissue extract, freshly harvested cells, or lysates of cells
20 which have been incubated in cell culture, in the presence of a detectably labeled antibody capable of identifying NGPCR gene products or conserved variants or peptide fragments thereof, and detecting the bound antibody by any of a number of techniques well-known in the art.

25 The biological sample may be brought in contact with and immobilized onto a solid phase support or carrier such as nitrocellulose, or other solid support which is capable of immobilizing cells, cell particles or soluble proteins. The support may then be washed with suitable buffers followed by
30 treatment with the detectably labeled NGPCR antibody or NGPCR ligand fusion protein. The solid phase support may then be washed with the buffer a second time to remove unbound antibody or fusion protein. The amount of bound label on solid support

may then be detected by conventional means.

By "solid phase support or carrier" is intended any support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, 5 polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite. The nature of the carrier can be either soluble to some extent or insoluble for the purposes of the present invention. The support material may have virtually any possible structural configuration so long 10 as the coupled molecule is capable of binding to an antigen or antibody. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the surface may be flat such as a sheet, test strip, etc. Preferred 15 supports include polystyrene beads. Those skilled in the art will know many other suitable carriers for binding antibody or antigen, or will be able to ascertain the same by use of routine experimentation.

The binding activity of a given lot of NGPCR antibody or 20 NGPCR ligand fusion protein may be determined according to well known methods. Those skilled in the art will be able to determine operative and optimal assay conditions for each determination by employing routine experimentation.

With respect to antibodies, one of the ways in which the 25 NGPCR antibody can be detectably labeled is by linking the same to an enzyme and use in an enzyme immunoassay (EIA) (Voller, A., "The Enzyme Linked Immunosorbent Assay (ELISA)", 1978, Diagnostic Horizons 2:1-7, Microbiological Associates Quarterly Publication, Walkersville, MD); Voller, A. et al., 1978, J. Clin. Pathol. 31:507-520; Butler, J.E., 1981, Meth. Enzymol. 73:482-523; 30 Maggio, E. (ed.), 1980, Enzyme Immunoassay, CRC Press, Boca Raton, FL,; Ishikawa, E. et al., (eds.), 1981, Enzyme Immunoassay, Kigaku Shoin, Tokyo). The enzyme that is bound to

the antibody will react with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety which can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Enzymes which can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods which employ a chromogenic substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

Detection may also be accomplished using any of a variety of other immunoassays. For example, by radioactively labeling the antibodies or antibody fragments, it is possible to detect NGPCR through the use of a radioimmunoassay (RIA) (see, for example, Weintraub, B., Principles of Radioimmunoassays, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, March, 1986, which is incorporated by reference herein). The radioactive isotope can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

It is also possible to label the antibody with a fluorescent compound. When the fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde and fluorescamine.

The antibody can also be detectably labeled using

fluorescence emitting metals such as ¹⁵²Eu, or others of the lanthanide series. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA).

5 The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged antibody is then determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful
10 chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester.

 Likewise, a bioluminescent compound may be used to label the antibody of the present invention. Bioluminescence is a type of
15 chemiluminescence found in biological systems in, which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are
20 luciferin, luciferase and aequorin.

5.5 SCREENING ASSAYS FOR COMPOUNDS THAT MODULATE NGPCR EXPRESSION OR ACTIVITY

 The following assays are designed to identify compounds that interact with (e.g., bind to) NGPCR (including, but not limited
25 to an ECD or CD of NGPCR), compounds that interact with (e.g., bind to) intracellular proteins that interact with NGPCR (including but not limited to the TM and CD of NGPCR), compounds that interfere with the interaction of NGPCR with transmembrane or intracellular proteins involved in NGPCR-mediated signal
30 transduction, and to compounds which modulate the activity of the NGPCR gene (i.e., modulate the level of NGPCR gene expression) or modulate the levels of NGPCR. Assays may additionally be

utilized which identify compounds that bind to NGPCR gene regulatory sequences (e.g., promoter sequences) and which may modulate NGPCR gene expression. See e.g., Platt, K.A., 1994, J. Biol. Chem. 269:28558-28562, which is incorporated herein by
5 reference in its entirety.

The compounds which may be screened in accordance with the invention include but are not limited to peptides, antibodies and fragments thereof, and other organic compounds (e.g., peptidomimetics) that bind to an ECD of the described NGPCR and
10 either mimic the activity triggered by the natural ligand (i.e., agonists) or inhibit the activity triggered by the natural ligand (i.e., antagonists); as well as peptides, antibodies or fragments thereof, and other organic compounds that mimic an ECD of the NGPCR (or a portion thereof) and bind to and "neutralize" natural
15 ligand.

Such compounds may include, but are not limited to, peptides such as, for example, soluble peptides, including but not limited to members of random peptide libraries; (see, e.g., Lam, K.S. et al., 1991, Nature 354:82-84; Houghten, R. et al., 1991, Nature
20 354:84-86), and combinatorial chemistry-derived molecular library made of D- and/or L- configuration amino acids, phosphopeptides (including, but not limited to members of random or partially degenerate, directed phosphopeptide libraries; see, e.g., Songyang, Z. et al., 1993, Cell 72:767-778), antibodies
25 (including, but not limited to, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and FAb, F(ab')₂ and FAb expression library fragments, and epitope-binding fragments thereof), and small organic or inorganic molecules.

30 Other compounds which can be screened in accordance with the invention include but are not limited to small organic molecules that are able to cross the blood-brain barrier, gain entry into

an appropriate cell (e.g., in the choroid plexus, the hypothalamus, etc.) and affect the expression of the NGPCR gene or some other gene involved in the NGPCR signal transduction pathway (e.g., by interacting with the regulatory region or transcription factors involved in gene expression); or such compounds that affect the activity of the NGPCR (e.g., by inhibiting or enhancing the enzymatic activity of a CD) or the activity of some other intracellular factor involved in the NGPCR signal transduction pathway.

10 Computer modeling and searching technologies permit identification of compounds, or the improvement of already identified compounds, that can modulate NGPCR expression or activity. Having identified such a compound or composition, the active sites or regions are identified. Such active sites might
15 typically be ligand binding sites. The active site can be identified using methods known in the art including, for example, from the amino acid sequences of peptides, from the nucleotide sequences of nucleic acids, or from study of complexes of the relevant compound or composition with its natural ligand. In the
20 latter case, chemical or X-ray crystallographic methods can be used to find the active site by finding where on the factor the complexed ligand is found.

Next, the three dimensional geometric structure of the active site is determined. This can be done by known methods,
25 including X-ray crystallography, which can determine a complete molecular structure. On the other hand, solid or liquid phase NMR can be used to determine certain intra-molecular distances. Any other experimental method of structure determination can be used to obtain partial or complete geometric structures. The
30 geometric structures may be measured with a complexed ligand, natural or artificial, which may increase the accuracy of the active site structure determined.

If an incomplete or insufficiently accurate structure is

determined, the methods of computer based numerical modeling can be used to complete the structure or improve its accuracy. Any recognized modeling method may be used, including parameterized models specific to particular biopolymers such as proteins or
5 nucleic acids, molecular dynamics models based on computing molecular motions, statistical mechanics models based on thermal ensembles, or combined models. For most types of models, standard molecular force fields, representing the forces between constituent atoms and groups, are necessary, and can be selected
10 from force fields known in physical chemistry. The incomplete or less accurate experimental structures can serve as constraints on the complete and more accurate structures computed by these modeling methods.

Finally, having determined the structure of the active site,
15 either experimentally, by modeling, or by a combination, candidate modulating compounds can be identified by searching databases containing compounds along with information on their molecular structure. Such a search seeks compounds having structures that match the determined active site structure and
20 that interact with the groups defining the active site. Such a search can be manual, but is preferably computer assisted. These compounds found from this search are potential NGPCR modulating compounds.

Alternatively, these methods can be used to identify
25 improved modulating compounds from an already known modulating compound or ligand. The composition of the known compound can be modified and the structural effects of modification can be determined using the experimental and computer modeling methods described above applied to the new composition. The altered
30 structure is then compared to the active site structure of the compound to determine if an improved fit or interaction results. In this manner systematic variations in composition, such as by varying side groups, can be quickly evaluated to obtain modified

modulating compounds or ligands of improved specificity or activity.

Further experimental and computer modeling methods useful to identify modulating compounds based upon identification of the active sites of the NGPCR, and related transduction and transcription factors will be apparent to those of skill in the art.

Examples of molecular modeling systems are the CHARMM and QUANTA programs (Polygen Corporation, Waltham, MA). CHARMM performs the energy minimization and molecular dynamics functions. QUANTA performs the construction, graphic modeling and analysis of molecular structure. QUANTA allows interactive construction, modification, visualization, and analysis of the behavior of molecules with each other.

A number of articles review computer modeling of drugs interactive with specific proteins, such as Rotivinen, *et al.*, 1988, *Acta Pharmaceutical Fennica* 97:159-166; Ripka, *New Scientist* 54-57 (June 16, 1988); McKinaly and Rossmann, 1989, *Annu. Rev. Pharmacol. Toxiciol.* 29:111-122; Perry and Davies, OSAR: Quantitative Structure-Activity Relationships in Drug Design pp. 189-193 (Alan R. Liss, Inc. 1989); Lewis and Dean, 1989 *Proc. R. Soc. Lond.* 236:125-140 and 141-162; and, with respect to a model receptor for nucleic acid components, Askew, *et al.*, 1989, *J. Am. Chem. Soc.* 111:1082-1090. Other computer programs that screen and graphically depict chemicals are available from companies such as BioDesign, Inc. (Pasadena, CA.), Allelix, Inc. (Mississauga, Ontario, Canada), and Hypercube, Inc. (Cambridge, Ontario). Although these are primarily designed for application to drugs specific to particular proteins, they can be adapted to design of drugs specific to regions of DNA or RNA, once that region is identified.

Although described above with reference to design and generation of compounds which could alter binding, one could also

screen libraries of known compounds, including natural products or synthetic chemicals, and biologically active materials, including proteins, for compounds which are inhibitors or activators.

5 Cell-based systems can also be used to identify compounds that bind the described NGPCR as well as assess the altered activity associated with such binding in living cells. One tool of particular interest for such assays is green fluorescent protein which is described, *inter alia*, in U.S. Patent No.
10 5,625,048, herein incorporated by reference. Cells that may be used in such cellular assays include, but are not limited to, leukocytes, or cell lines derived from leukocytes, lymphocytes, stem cells, including embryonic stem cells, and the like. In addition, expression host cells (e.g., B95 cells, COS cells, CHO
15 cells, OMK cells, fibroblasts, Sf9 cells) genetically engineered to express functional NGPCR and to respond to activation by the test, or natural, ligand, as measured by a chemical or phenotypic change, or induction of another host cell gene, can be used as an end point in the assay.

20 Compounds identified via assays such as those described herein may be useful, for example, in elaborating the biological function of a NGPCR gene product. Such compounds can be administered to a patient at therapeutically effective doses to treat any of a variety of physiological or mental disorders. A
25 therapeutically effective dose refers to that amount of the compound sufficient to result in any amelioration, impediment, prevention, or alteration of any biological or overt symptom.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures
30 or experimental animals, e.g., for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic

index and it can be expressed as the ratio LD_{50}/ED_{50} . Compounds which exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

Pharmaceutical compositions for use in accordance with the present invention may be formulated in conventional manner using one or more physiologically acceptable carriers or excipients. Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by inhalation or insufflation (either through the mouth or the nose) or oral, buccal, parenteral, intracranial, intrathecal, or rectal administration.

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by

conventional means with pharmaceutically acceptable excipients such as binding agents (e.g., pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (e.g., sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (e.g., lecithin or acacia); non-aqueous vehicles (e.g., almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (e.g., methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate.

Preparations for oral administration may be suitably formulated to give controlled release of the active compound.

For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be

determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g. gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

5 The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multi-dose containers, with an added preservative. The compositions may
10 take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

15 The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides.

 In addition to the formulations described previously, the
20 compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as
25 an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

 The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms
30 containing the active ingredient. The pack may for example comprise metal or plastic foil; such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

5.5.1 IN VITRO SCREENING ASSAYS FOR COMPOUNDS THAT BIND TO NGPCR

In vitro systems may be designed to identify compounds capable of interacting with (e.g., binding to) the described NGPCR (including, but not limited to, a ECD or CD of NGPCR).

- 5 Compounds identified may be useful, for example, in modulating the activity of wild type and/or mutant NGPCR gene products; may be useful in elaborating the biological function of the NGPCR; may be utilized in screens for identifying compounds that disrupt normal NGPCR interactions; or may in themselves disrupt such
- 10 interactions.

The principle of the assays used to identify compounds that bind to the NGPCR involves preparing a reaction mixture of the NGPCR and the test compound under conditions and for a time sufficient to allow the two components to interact and bind, thus

15 forming a complex which can be removed and/or detected in the reaction mixture. The NGPCR species used can vary depending upon the goal of the screening assay. For example, where agonists of the natural ligand are sought, the full length NGPCR, or a soluble truncated NGPCR, e.g., in which the TM and/or CD is

20 deleted from the molecule, a peptide corresponding to a ECD or a fusion protein containing one or more NGPCR ECD fused to a protein or polypeptide that affords advantages in the assay system (e.g., labeling, isolation of the resulting complex, etc.) can be utilized. Where compounds that interact with the

25 cytoplasmic domain are sought to be identified, peptides corresponding to the NGPCR CD and fusion proteins containing the NGPCR CD can be used.

The screening assays can be conducted in a variety of ways. For example, one method to conduct such an assay would involve

30 anchoring the NGPCR protein, polypeptide, peptide or fusion protein or the test substance onto a solid phase and detecting NGPCR/test compound complexes anchored on the solid phase at the

end of the reaction. In one embodiment of such a method, the NGPCR reactant may be anchored onto a solid surface, and the test compound, which is not anchored, may be labeled, either directly or indirectly.

5 In practice, microtiter plates may conveniently be utilized as the solid phase. The anchored component may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished by simply coating the solid surface with a solution of the protein and drying. Alternatively, an
10 immobilized antibody, preferably a monoclonal antibody, specific for the protein to be immobilized may be used to anchor the protein to the solid surface. The surfaces may be prepared in advance and stored.

In order to conduct the assay, the nonimmobilized component
15 is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (e.g., by washing) under conditions such that any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be
20 accomplished in a number of ways. Where the previously nonimmobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously nonimmobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the
25 surface; e.g., using a labeled antibody specific for the previously nonimmobilized component (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody).

Alternatively, a reaction can be conducted in a liquid
30 phase, the reaction products separated from unreacted components, and complexes detected; e.g., using an immobilized antibody specific for the NGPCR protein, polypeptide, peptide or fusion protein or the test compound to anchor any complexes formed in

solution, and a labeled antibody specific for the other component of the possible complex to detect anchored complexes.

Alternatively, cell-based assays can be used to identify compounds that interact with NGPCR. To this end, cell lines that
5 express NGPCR, or cell lines (e.g., COS cells, CHO cells, fibroblasts, etc.) that have been genetically engineered to express NGPCR (e.g., by transfection or transduction of NGPCR DNA) can be used. Interaction of the test compound with, for example, a ECD from NGPCR expressed by the host cell can be
10 determined by comparison or competition with native ligand.

5.5.2 ASSAYS FOR INTRACELLULAR PROTEINS THAT INTERACT WITH NGPCR

Any method suitable for detecting protein-protein interactions may be employed for identifying transmembrane proteins or intracellular proteins that interact with the
15 described NGPCR. Among the traditional methods which may be employed are co-immunoprecipitation, crosslinking and co-purification through gradients or chromatographic columns of cell lysates or proteins obtained from cell lysates to identify proteins in the lysate that interact with NGPCR. For these
20 assays, the NGPCR component used can be a full length NGPCR, a soluble derivative lacking the membrane-anchoring region (e.g., a truncated NGPCR in which a TM is deleted resulting in a truncated molecule containing a ECD fused to a CD), a peptide corresponding to a CD or a fusion protein containing a CD from the described
25 NGPCR. Once isolated, such an intracellular protein can be identified and can, in turn, be used, in conjunction with standard techniques, to identify proteins with which it interacts. For example, at least a portion of the amino acid sequence of an intracellular protein which interacts with NGPCR
30 can be ascertained using techniques well known to those of skill in the art, such as via the Edman degradation technique. (See,

e.g., Creighton, 1983, "Proteins: Structures and Molecular Principles", W.H. Freeman & Co., N.Y., pp.34-49). The amino acid sequence obtained may be used as a guide for the generation of oligonucleotide mixtures that can be used to screen for gene sequences encoding such intracellular proteins. Screening may be accomplished, for example, by standard hybridization or PCR techniques. Techniques for the generation of oligonucleotide mixtures and the screening are well-known. (See, e.g., Ausubel, *supra*, and PCR Protocols: A Guide to Methods and Applications, 1990, Innis, M. et al., eds. Academic Press, Inc., New York).

Additionally, methods may be employed which result in the simultaneous identification of genes which encode the transmembrane or intracellular proteins interacting with NGPCR. These methods include, for example, probing expression libraries, in a manner similar to the well known technique of antibody probing of λ gt11 libraries, using labeled NGPCR protein, or an NGPCR polypeptide, peptide or fusion protein, e.g., an NGPCR polypeptide or NGPCR domain fused to a marker (e.g., an enzyme, fluor, luminescent protein, or dye), or an Ig-Fc domain.

One method that detects protein interactions *in vivo*, the two-hybrid system, is described in detail for illustration only and not by way of limitation. One version of this system has been described (Chien et al., 1991, Proc. Natl. Acad. Sci. USA, 88:9578-9582) and is commercially available from Clontech (Palo Alto, CA).

Briefly, utilizing such a system, plasmids are constructed that encode two hybrid proteins: one plasmid consists of nucleotides encoding the DNA-binding domain of a transcription activator protein fused to NGPCR nucleotide sequence encoding NGPCR, or a NGPCR polypeptide, peptide or fusion protein, and the other plasmid consists of nucleotides encoding the transcription activator protein's activation domain fused to a cDNA encoding an

unknown protein which has been recombined into this plasmid as part of a cDNA library. The DNA-binding domain fusion plasmid and the cDNA library are transformed into a strain of the yeast *Saccharomyces cerevisiae* that contains a reporter gene (e.g., HBS or *lacZ*) whose regulatory region contains the transcription activator's binding site. Either hybrid protein alone cannot activate transcription of the reporter gene: the DNA-binding domain hybrid cannot because it does not provide activation function and the activation domain hybrid cannot because it cannot localize to the activator's binding sites. Interaction of the two hybrid proteins reconstitutes the functional activator protein and results in expression of the reporter gene, which is detected by an assay for the reporter gene product.

The two-hybrid system or related methodology may be used to screen activation domain libraries for proteins that interact with the "bait" gene product. By way of example, and not by way of limitation, the NGPCR may be used as the bait gene product. Total genomic or cDNA sequences are fused to the DNA encoding an activation domain. This library and a plasmid encoding a hybrid of the bait NGPCR gene product fused to the DNA-binding domain are cotransformed into a yeast reporter strain, and the resulting transformants are screened for those that express the reporter gene. For example, and not by way of limitation, a bait NGPCR gene sequence, such as the open reading frame of NGPCR (or a domain of NGPCR) can be cloned into a vector such that it is translationally fused to the DNA encoding the DNA-binding domain of the GAL4 protein. These colonies are purified and the library plasmids responsible for reporter gene expression are isolated. DNA sequencing is then used to identify the proteins encoded by the library plasmids.

A cDNA library of the cell line from which proteins that interact with bait NGPCR gene product are to be detected can be made using methods routinely practiced in the art. According to

the particular system described herein, for example, the cDNA fragments can be inserted into a vector such that they are translationally fused to the transcriptional activation domain of GAL4. This library can be co-transformed along with the bait NGPCR gene-GAL4 fusion plasmid into a yeast strain which contains a lacZ gene driven by a promoter which contains GAL4 activation sequence. A cDNA encoded protein, fused to GAL4 transcriptional activation domain, that interacts with bait NGPCR gene product will reconstitute an active GAL4 protein and thereby drive expression of the HIS3 gene. Colonies that express HIS3 can be detected by their growth on petri dishes containing semi-solid agar based media lacking histidine. The cDNA can then be purified from these strains, and used to produce and isolate the bait NGPCR gene-interacting protein using techniques routinely practiced in the art.

5.5.3 ASSAYS FOR COMPOUNDS THAT INTERFERE WITH NGPCR/INTRACELLULAR OR NGPCR/TRANSMEMBRANE MACROMOLECULE INTERACTION

The macromolecules that interact with the NGPCR are referred to, for purposes of this discussion, as "binding partners". These binding partners are likely to be involved in the NGPCR signal transduction pathway. Therefore, it is desirable to identify compounds that interfere with or disrupt the interaction of such binding partners which may be useful in regulating the activity of NGPCR and controlling disorders associated with NGPCR activity.

The basic principle of the assay systems used to identify compounds that interfere with the interaction between NGPCR and its binding partner or partners involves preparing a reaction mixture containing NGPCR protein, polypeptide, peptide or fusion protein as described in Sections 5.5.1 and 5.5.2 above, and the binding partner under conditions and for a time sufficient to allow the two to interact and bind, thus forming a complex. In

order to test a compound for inhibitory activity, the reaction mixture is prepared in the presence and absence of the test compound. The test compound may be initially included in the reaction mixture, or may be added at a time subsequent to the addition of the NGPCR moiety and its binding partner. Control reaction mixtures are incubated without the test compound or with a placebo. The formation of any complexes between the NGPCR moiety and the binding partner is then detected. The formation of a complex in the control reaction, but not in the reaction mixture containing the test compound, indicates that the compound interferes with the interaction of the NGPCR and the interactive binding partner. Additionally, complex formation within reaction mixtures containing the test compound and normal NGPCR protein may also be compared to complex formation within reaction mixtures containing the test compound and a mutant NGPCR. This comparison may be important in those cases wherein it is desirable to identify compounds that specifically disrupt interactions of mutant, or mutated, NGPCRs but not normal NGPCRs.

The assay for compounds that interfere with the interaction of the described NGPCR and binding partners can be conducted in a heterogeneous or homogeneous format. Heterogeneous assays involve anchoring either the NGPCR moiety product or the binding partner onto a solid phase and detecting complexes anchored on the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction by competition can be identified by conducting the reaction in the presence of the test substance; *i.e.*, by adding the test substance to the reaction mixture prior to, or simultaneously with, the NGPCR moiety and interactive binding partner. Alternatively, test compounds that disrupt preformed complexes,

e.g. compounds with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed. The various formats are described briefly below.

5 In a heterogeneous assay system, either the NGPCR moiety or an interactive binding partner, is anchored onto a solid surface, while the non-anchored species is labeled, either directly or indirectly. In practice, microtiter plates are conveniently utilized. The anchored species may be immobilized by non-
10 covalent or covalent attachments. Non-covalent attachment may be accomplished simply by coating the solid surface with a solution of the NGPCR gene product or binding partner and drying. Alternatively, an immobilized antibody specific for the species to be anchored may be used to anchor the species to the solid
15 surface. The surfaces may be prepared in advance and stored.

 In order to conduct the assay, the partner of the immobilized species is exposed to the coated surface with or without the test compound. After the reaction is complete, unreacted components are removed (e.g., by washing) and any
20 complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the non-
25 immobilized species is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific for the initially non-immobilized species (the antibody, in turn, may be directly labeled or indirectly labeled with a labeled anti-Ig antibody). Depending
30 upon the order of addition of reaction components, test compounds which inhibit complex formation or which disrupt preformed complexes can be detected.

 Alternatively, the reaction can be conducted in a liquid

phase in the presence or absence of the test compound, the reaction products separated from unreacted components, and complexes detected; e.g., using an immobilized antibody specific for one of the binding components to anchor any complexes formed in solution, and a labeled antibody specific for the other partner to detect anchored complexes. Again, depending upon the order of addition of reactants to the liquid phase, test compounds which inhibit complex or which disrupt preformed complexes can be identified.

10 In an alternate embodiment of the invention, a homogeneous assay can be used. In this approach, a preformed complex of the described NGPCR moiety and an interactive binding partner is prepared in which either the NGPCR or its binding partners is labeled, but the signal generated by the label is quenched due to formation of the complex (see, e.g., U.S. Patent No. 4,109,496 by Rubenstein which utilizes this approach for immunoassays). The addition of a test substance that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above background. In this way, test substances which disrupt NGPCR/intracellular binding partner interaction can be identified.

In a particular embodiment, a NGPCR fusion can be prepared for immobilization. For example, the NGPCR or a peptide fragment, e.g., corresponding to a CD, can be fused to a glutathione-S-transferase (GST) gene using a fusion vector, such as pGEX-5X-1, in such a manner that its binding activity is maintained in the resulting fusion protein. The interactive binding partner can be purified and used to raise a monoclonal antibody, using methods routinely practiced in the art and described above, in Section 5.3. This antibody can be labeled with the radioactive isotope ^{125}I , for example, by methods routinely practiced in the art. In a heterogeneous assay, e.g., the GST-NGPCR fusion protein can be anchored to glutathione-

agarose beads. The interactive binding partner can then be added in the presence or absence of the test compound in a manner that allows interaction and binding to occur. At the end of the reaction period, unbound material can be washed away, and the
5 labeled monoclonal antibody can be added to the system and allowed to bind to the complexed components. The interaction between the NGPCR gene product and the interactive binding partner can be detected by measuring the amount of radioactivity that remains associated with the glutathione-agarose beads. A
10 successful inhibition of the interaction by the test compound will result in a decrease in measured radioactivity.

Alternatively, the GST-NGPCR fusion protein and the interactive binding partner can be mixed together in liquid in the absence of the solid glutathione-agarose beads. The test
15 compound can be added either during or after the species are allowed to interact. This mixture can then be added to the glutathione-agarose beads and unbound material is washed away. Again the extent of inhibition of the NGPCR/binding partner interaction can be detected by adding the labeled antibody and
20 measuring the radioactivity associated with the beads.

In another embodiment of the invention, these same techniques can be employed using peptide fragments that correspond to the binding domains of NGPCR and/or the interactive or binding partner (in cases where the binding partner is a
25 protein), in place of one or both of the full length proteins. Any number of methods routinely practiced in the art can be used to identify and isolate the binding sites. These methods include, but are not limited to, mutagenesis of the gene encoding one of the proteins and screening for disruption of binding in a
30 co-immunoprecipitation assay. Compensatory mutations in the gene encoding the second species in the complex can then be selected. Sequence analysis of the genes encoding the respective proteins will reveal the mutations that correspond to the region of the

protein involved in interactive binding. Alternatively, one protein can be anchored to a solid surface using methods described above, and allowed to interact with and bind to its labeled binding partner, which has been treated with a
5 proteolytic enzyme, such as trypsin. After washing, a relatively short, labeled peptide comprising the binding domain may remain associated with the solid material, which can be isolated and identified by amino acid sequencing. Also, once the gene coding for the intracellular binding partner is obtained, short gene
10 segments can be engineered to express peptide fragments of the protein, which can then be tested for binding activity and purified or synthesized.

For example, and not by way of limitation, the NGPCR gene product can be anchored to a solid material as described, above,
15 by making a GST-NGPCR fusion protein and allowing it to bind to glutathione agarose beads. The interactive binding partner can be labeled with a radioactive isotope, such as ^{35}S , and cleaved with a proteolytic enzyme such as trypsin. Cleavage products can then be added to the anchored GST-NGPCR fusion protein and
20 allowed to bind. After washing away unbound peptides, labeled bound material, representing the intracellular binding partner binding domain, can be eluted, purified, and analyzed for amino acid sequence by well-known methods. Peptides so identified can be produced synthetically or fused to appropriate facilitative
25 proteins using recombinant DNA technology.

The present invention is not to be limited in scope by the specific embodiments described herein, which are intended as single illustrations of individual aspects of the invention, and functionally equivalent methods and components are within the
30 scope of the invention. Indeed, various modifications of the invention, in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and accompanying drawings. Such modifications are

intended to fall within the scope of the appended claims.

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2